Title page

Title: Predicting age using neuroimaging: innovative brain ageing biomarkers

Authors:

James H Cole PhD, Computational, Cognitive & Clinical Neuroimaging Laboratory, Imperial College London, UK.

Katja Franke PhD, Structural Brain Mapping Group, University Hospital Jena, Germany.

Correspondence: james.cole@imperial.ac.uk; katja.franke@uni-jena.de

Keywords (two to six):

Brain ageing; Neuroimaging; ageing biomarker; brain diseases; machine learning

Counts:

Abstract: 120/120

Main text: 2788/2500

References: 61/60

Abstract

The brain changes as we age and these changes are associated with functional deterioration and neurodegenerative disease. It is vital that we better understand individual differences in the brain ageing process, hence techniques for making individualised predictions of brain ageing have been developed. We present evidence supporting use of neuroimaging-based *brain age* as a biomarker of an individual's brain health. Increasingly, research is showing how brain diseases or poor physical health negatively impact *brain age*. Importantly, recent evidence shows that having an 'older'-appearing brain relates to advanced physiological and cognitive ageing and the risk of mortality. We discuss controversies surrounding *brain age* and highlight emerging trends such as the use of multi-modality neuroimaging and the employment of 'deep learning' methods.

Brain scans can be used to predict age

As the global population ages, the burden of age-associated functional decline and disease is increasing [1]. Methods are required for predicting who is at higher risk of age-associated deterioration, how this decline will progress and which treatments are most appropriate. The ageing process is biologically complex [2], and despite the generally negative effects of ageing, there is pronounced variation among people in the timing of manifestation of ageing effects (Figure 1). This variation in brain aging may contribute to the enormous variation in human lifespan, and in the varying ages at which people develop age-related diseases. Potentially, a person's underlying **biological age** (see Glossary) may differ from their chronological age, and could be a better indicator of future risk of experiencing age-associated health issues.

Ageing results in marked changes to the structure and function of the brain. Cognitive decline and an increased risk of neurodegenerative diseases are a key source of the burden caused by ageing. However, pronounced individual differences are also seen in measures of the brain as people age [3]. While the average age-driven trajectories of brain volume, cortical thickness and white matter microstructure have been characterised in healthy people [4-6], a single person may differ considerably from the average. Potentially, the extent to which someone deviates from healthy brain ageing trajectories could indicate underlying problems in outwardly healthy people, and relate to the risk of cognitive ageing or age-associated brain disease. Hence, reliable biomarkers of brain ageing could be of great neuroscientific and clinical value.

Using structural or functional neuroimaging data, it is now possible to predict age [7, 8]. The most effective approaches to age prediction have used data from **magnetic resonance imaging (MRI)** scans of the brain and run a type of statistical analysis on the images called **machine learning**. By 'learning' the relationship between patterns of data from brain scans and chronological age in a training dataset of healthy people, age predictions can be made using brain images from people not included in the initial training. The most accurate measures in adults have reported a mean absolute error (MAE) of <5 years [8-12], which can be measured with high test-retest reliability [9, 13]. Moreover, in studies covering age ranges between early childhood and young adulthood the most accurate predictions result in MAEs of approximately only 1 year [14-16].

While using neuroimaging to predict age may be seen as an interesting academic exercise, it is

also an important proof-of-concept, showing that the information extracted from a single MRI scan relates strongly to chronological age and that it can be used to make accurate age estimations in new scans. Furthermore, a growing body of research is demonstrating that so-called ***brain age*** has both clinical and broader scientific relevance. This paradigm has provided a new way to explore how the brain changes during ageing and how brain diseases interact with 'normal' brain ageing. Potentially, *brain age* could be used as a personalised biomarker of brain health during ageing, and this individual-specific nature is particularly important. The extensive study of group-mean differences in case-control studies of brain diseases has yielded few clinical applications. Conversely, *brain age* locates an individual within a normative ageing distribution. If this location can be shown to be relevant for health outcomes, then *brain age* presents a framework for applying neuroimaging clinically to characterise brain health. Here we outline the methods for predicting *brain age*, evaluate evidence for its use as an ageing biomarker and discuss trends in the ongoing development of the paradigm.

How does neuroimaging-based *brain age* prediction work?

The accuracy of *brain age* prediction relies on the fact that the brain changes as we age (Box 1) and that these changes are reasonably consistent between different people. Neuroimaging provides a unique window into the brain ageing process, allowing precise and reliable measurement of many aspects of brain structure and function. Recent advances in computing and the increasing availability of large neuroimaging datasets mean that researchers have been enabled to apply machine learning to the problem of age prediction (Figure 2, Key Figure; Box 2).

How does *brain age* relate to other ageing measures?

The brain can be affected by peripheral physiological changes and having a healthy brain is essential for overall health. Therefore, measuring *brain age* could provide a window for general biological ageing, as a potential **ageing biomarker**. To that end, it is useful to consider whether *brain age* relates to other known facets of ageing. Measures of ageing typically used in **gerontology** include physiological, cognitive and biological components. Physiological measurements of hand-grip strength, lung function, and walking speed are used to characterise general physical health as well as to predict risk of mortality in older adults [17-19]. Evidently, by using robust measurements techniques as proxies of underlying physiological variability,

information about general health and residual lifespan can be obtained. *Brain age* appears to meet these criteria, as a recent large-scale study in 73 year-olds found a significant relationship between *brain age* and mortality risk, ascertained up to seven years after scanning [20]. For every year that an individual's brain was predicted to be older than their chronological age, there was a 6% increased risk of death. This study also showed that lower grip-strength, lower forced expiratory volume and a slower walking time were all significantly associated with *brain age*, as was a composite measure of fluid cognition. This supports the idea that the brain is sensitive to general declines in health and suggests that *brain age* could be used as an ageing biomarker, to make individualised predictions about mortality risk in older adults. The abovementioned study also compared *brain age* to other putative ageing biomarkers. *Brain age* did not correlate with either leukocyte telomere length or *DNA-methylation age*. Interestingly, *brain age* was a stronger predictor of mortality than the other measures, though a combined model of *brain age* with *DNA-methylation age* was the best predictor, illustrating the benefits of combining distinct ageing biomarkers.

Brain diseases and *brain age*

If *brain age* can provide information about future health outcomes in the general population, this motivates research into potential causes of deviations from healthy brain ageing. As such, considering how specific diseases relate to *brain age* may help isolate deleterious influences on brain health in later life. Understanding how variability in *brain age* within diverse clinical samples relates to other facets of non-communicable and age-related diseases could help identify individuals at risk of poor health outcomes as ageing and disease processes interact. Potentially, diseases result in increases to brain ageing, as a one-off 'hit' or a progressive acceleration to the process. Alternatively, the presence of a disease may not cause brain ageing *per se*, but occurs on top of underlying individual differences in 'normal' brain ageing. This could mean that the effects of that disease are exacerbated in those with 'older' as opposed to 'younger' appearing brains. Either way, measuring *brain age* in disease groups could be fruitful for quantifying some of the heterogeneity within a disease, improving identification of individuals at higher risk of poor health outcomes. Consequently, *brain age* could be used as general marker of poorer brain health to help stratify enrolment of individuals into clinical trials of therapies aimed at improving brain health in older adults, who may not have observable clinical or cognitive impairments.

While aetiologically and pathophysiologically distinct, many diseases seem to have common, secondary effects on the brain. For example, brain injury, multiple sclerosis, major depressive disorder and Alzheimer's are all associated with a heightened immune response, neuroinflammation, oxidative stress, mitochondrial dysfunction and epigenetic alterations [21-30]. Notably, all these phenomena are also implicated in the biology of 'normal' ageing [2]. Furthermore, a number of diseases have been proposed to exacerbate biological ageing, including Down's syndrome, HIV and traumatic brain injury [31-33]. Given the relationship between ageing and disease risk, it is unsurprising that common underlying mechanisms may be present. However, the availability of ageing biomarkers now allows researchers to evaluate evidence of abnormal ageing in specific diseases, and in the context of brain diseases, *brain age* is likely to be a particularly relevant measure. It is hoped that combining ageing-related biomarkers with more disease-specific biomarkers will lead to further improvements in diagnostic and prognostic modelling, moving closer to clinical applications of neuroimaging.

Indeed, a growing number of neuropsychiatric diseases have been associated with increases in *brain age* (Table 1). These include traumatic brain injury [34], schizophrenia [12, 35, 36], epilepsy [37], Down's syndrome [38], HIV [39], mild cognitive impairment and Alzheimer's [13, 40-42]. Similar results have also been seen in peripheral conditions and non-communicable diseases, such as mid-life obesity [43] and diabetes [44], suggesting again that the brain is also sensitive to deteriorations in general physical health. These outwardly disparate conditions may share some common pathological neurobiological components, effects secondary to the disparate primary pathological processes, that result in an increase in age-associated loss of brain volume. Interestingly, *brain age* was more sensitive in showing differences between groups, when compared to total and regional brain volumes [40]. Methodologically, the variance in *brain age* is largely explained by a composite of brain volume, age and sex, though also contains unique variance not captured by commonly used measures. Thus, analysing *brain age* in these contexts provides a novel way to capture individual differences within the general population as well as disease groups, which relate to additional facets of various diseases or even predict future outcomes. For example, increased *brain age* in people with mild cognitive impairment has been associated with greater risk of developing Alzheimer's within three years [40, 42].

Despite the many different causes of neuropathology, response mechanisms of the brain seem to be relatively limited, whether the causes are infectious, traumatic or genetic. Hence, the *brain*

age studies can be seen as evidence that common secondary mechanisms, observed across diseases, may relate to those seen in healthy ageing and may be important for some of the neurological, cognitive and behavioural consequences of brain diseases. In line with this, cognitive performance has also been assessed in studies of *brain age*. In general, there are significant relationships between global cognitive performance and *brain age*, being more pronounced in disease samples [10, 13, 15, 20, 34, 39, 40, 42]. This supports the idea that brain ageing and cognitive ageing are linked, though the modest strength of these associations suggests that further methods development is needed to better capture the variation in brain structure and cognitive performance.

Improving individual brain health

While there may be many deleterious influences on *brain age*, there is also evidence of protective factors. Significant associations with decreased *brain age* and markers of good health in cognitively healthy elderly [45] and the general population [46] have been reported. Furthermore, the number of years of education and a self-reported measure of physical activity (number of stairs climbed daily) were reported to be significantly associated with a lower *brain age* in individuals aged 19-79 years [47]. Alongside this, recent studies have observed a reduction in *brain age* in long-term practitioners of meditation [48] and in amateur musicians [49].

Although only cross-sectional, such results are promising. They suggest that interventions could be effective in slowing or potentially even reversing brain ageing, reducing the risk of future cognitive decline and age-associated disease. However, prospective longitudinal studies of positive influences on *brain age* are yet to be conducted. This represents the next, and crucial, step in developing a framework for evaluating potential treatments of age-associated brain deterioration.

Controversies around *brain age*

While the *brain age* paradigm offers a powerful approach to investigating brain ageing, it has attracted some criticisms, either technical or philosophical. For instance, some consider the only factor that affects age to be time, thus ageing *per se* cannot deviate from its chronological course. This criticism applies to all potential ageing biomarkers, instead suggesting that in fact there is limited biological variability in ageing and that deviations are due to specific pathological processes, not reflecting an extension of ‘normal’ ageing. However, there is strong support for

the hypothesis that ageing results from cumulative biological damage [50]. It follows from this that variable exposure to the causes of this cumulative damage would result in individual differences in rates of underlying biological. Furthermore, the very fact that ageing is the major risk for numerous diseases strongly suggests that biological ageing and disease are intrinsically linked. Beyond this, we argue that whether or not an increased *brain age* indicates that a brain is actually 'older' is not the chief consideration. If *brain age* (or other biologically-predicted ages for that matter) can be a useful neuroscientific and clinical tool then it warrants further exploration.

Another criticism of *brain age* is that by condensing whole-brain **voxel**-wise information into a single number it is overly 'black box'. By not scrutinising exactly which **features** of a brain scan are used for predicting age, important neuroscientific information may be disregarded and it is unclear precisely what information age prediction is based on. However, there are several important reasons why interpreting the **'weight maps'** derived from machine learning is complicated and does not offer a straightforward interpretation in the context of brain ageing [51]. First, no one part of the brain is the sole driver of ageing; brain ageing is a global phenomenon. Second, age-related changes to the brain are subtle, non-linear, spatially distributed and vary between individuals [4, 6, 52]. The advantage of the *brain age* paradigm is that, by using machine learning, the model can learn a range of different brain structures that may be healthy. This avoids reductively focusing on the average, which likely is unrepresentative of any single individual.

A final criticism of *brain age* is that it relies on using the resulting error in prediction (i.e., the difference between the predicted age and chronological age) as a metric for further analysis. Statistically, this is equivalent to using the residuals for an individual from a linear regression model. Basing clinical or neuroscientific interpretations on error may be semantically dubious, as in theory more accurate models would reduce this error. Crucially, however, the key to determining the validity of *brain age* lies in external validation with other characteristics measured in the same individuals. For example, the fact that the error metric (e.g., brain-predicted age difference) relates to cognitive performance, ageing fitness and subsequent survival [13, 15, 20, 40, 42, 47], strongly supports the idea that by quantifying this error, clinically and biologically meaningful insights can be derived.

Concluding remarks

The emerging field of *brain age* prediction is evolving rapidly and an increasing number of researchers are employing *brain age* analysis to explore brain ageing in health and disease. A number of promising trends are developing. These include the combination of multiple neuroimaging modalities, for example combining structural and functional MRI data, or multiple structural MRI modalities (T2*, diffusion-MRI), resulting in improved prediction performance [10, 53]. Combined predictors potentially better capture the various facets of brain ageing, including brain atrophy, iron deposition and alterations to white matter microstructure (see Outstanding Questions).

Another development is the increasing availability of large datasets. Key to accurate machine learning is having a sufficient number of examples to learn from. Initiatives such as the International Neuroimaging Data-Sharing Initiative (INDI; http://fcon_1000.projects.nitrc.org/) and NeuGrid4U (https://neugrid4you.eu/) encourage sharing existing datasets. Ground-breaking projects like the Human Connectome Project and UK Biobank have been explicitly designed to share data and are making unprecedented amounts of neuroimaging data accessible.

Important for leveraging these larger and more complex datasets is innovation in computational statistics, to optimise algorithms for predicting *brain age* [54]. In particular, **deep learning** methods show particular promise [9]. The 'hidden' layers in deep learning allow data-driven representation of different global and local data features, meaning that hitherto unknown relationships can be more accurately identified. One benefit particular to neuroimaging is the removal the reliance on data pre-processing to extract meaningful features. Such features are automatically encoded by the deep neural network, avoiding the model-dependent decisions used in image pre-processing (e.g., registration algorithm, template selection). While the computational demands for deep learning are high, the added benefits likely outweigh the costs, and deep learning might enjoy increasing interest in *brain age* analysis, as in other neuroimaging research.

The ability to predict a person's *brain age* using neuroimaging data is increasingly providing insights into both positive and negative effects on age-associated brain changes, and is shedding new light on how diseases affect the ageing brain. Furthermore, *brain age* has the potential to identify individuals at risk of experiencing advanced biological ageing, and thus could

provide a biomarker of age-associated health problems. As the technical aspects of *brain age* analysis are further developed, the possibility that neuroimaging-based measures of *brain age* could be used to evaluate neuroprotective preventions and therapeutics comes closer to being realised.

References

1 Vos, T.*, et al.* (2012) Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* 380, 2163-2196

2 Lopez-Otin, C.*, et al.* (2013) The hallmarks of aging. *Cell* 153, 1194-1217

3 Raz, N.*, et al.* (2005) Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cereb Cort* 15, 1676-1689

4 Raz, N.*, et al.* (2010) Trajectories of brain aging in middle-aged and older adults: Regional and individual differences. *NeuroImage* 51, 501-511

5 Salat, D.H.*, et al.* (2005) Age-related alterations in white matter microstructure measured by diffusion tensor imaging. *Neurobiol Aging* 26, 1215-1227

6 Storsve, A.B.*, et al.* (2014) Differential longitudinal changes in cortical thickness, surface area and volume across the adult life span: Regions of accelerating and decelerating change. *J Neurosci* 34, 8488-8498

7 Dosenbach, N.U.F.*, et al.* (2010) Prediction of Individual Brain Maturity Using fMRI. *Science* 329, 1358-1361

8 Franke, K.*, et al.* (2010) Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. *NeuroImage* 50, 883-892

9 Cole, J.H.*, et al.* (2017) Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage* 163C, 115-124

10 Liem, F.*, et al.* (2017) Predicting brain-age from multimodal imaging data captures cognitive impairment. *NeuroImage* 148, 179-188

11 Lin, L.*, et al.* (2016) Predicting healthy older adult's brain age based on structural connectivity networks using artificial neural networks. *Computer methods and programs in biomedicine* 125, 8-17

12 Schnack, H.G.*, et al.* (2016) Accelerated Brain Aging in Schizophrenia: A Longitudinal Pattern Recognition Study. *Am J Psych* 173, 607-616

13 Franke, K. and Gaser, C. (2012) Longitudinal changes in individual BrainAGE in healthy aging, mild cognitive impairment, and Alzheimer's Disease. *GeroPsych* 25, 235-245

14 Brown, T.*, et al.* (2012) Neuroanatomical Assessment of Biological Maturity. *Curr Biol* 22, 1693-1698

15 Erus, G.*, et al.* (2015) Imaging Patterns of Brain Development and their Relationship to Cognition. *Cereb Cortex* 25, 1676-1684

16 Franke, K.*, et al.* (2012) Brain maturation: Predicting individual BrainAGE in children and adolescents using structural MRI. *NeuroImage* 63, 1305-1312

17 Leong, D.P.*, et al.* (2015) Prognostic value of grip strength: findings from the Prospective Urban Rural Epidemiology (PURE) study. *The Lancet 386*, 266–273

18 Studenski, S.*, et al.* (2011) Gait speed and survival in older adults. *JAMA* 305, 50-58

19 Schunemann, H.J.*, et al.* (2000) Pulmonary function is a long-term predictor of mortality in the general population: 29-year follow-up of the Buffalo Health Study. *Chest* 118, 656-664

20 Cole, J.H.*, et al.* (2017) Brain age predicts mortality. *Mol Psych*

21 Herrup, K. (2010) Reimagining Alzheimer's disease - An age-based hypothesis. *J Neurosci* 30, 16755-16762

22 Lin, M.T. and Beal, M.F. (2006) Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* 443, 787-795

23 Ramlackhansingh, A.F.*, et al.* (2011) Inflammation after trauma: Microglial activation and traumatic brain injury. *Ann Neurol* 70, 374-383

24 Lassmann, H.*, et al.* (2012) Progressive multiple sclerosis: Pathology and pathogenesis. *Nat Rev Neurol* 8, 647-656

25 Lu, F.*, et al.* (2000) Oxidative damage to mitochondrial DNA and activity of mitochondrial enzymes in chronic active lesions of multiple sclerosis. *Journal of the Neurological Sciences* 177, 95-103

26 Leonard, B. and Maes, M. (2012) Mechanistic explanations how cell-mediated immune activation, inflammation and oxidative and nitrosative stress pathways and their sequels and concomitants play a role in the pathophysiology of unipolar depression. *Neuroscience and Biobehavioral Reviews* 36, 764-785

27 Lardenoije, R.*, et al.* (2015) The epigenetics of aging and neurodegeneration. *Progress in Neurobiology* 131, 21-64

28 Koch, M.W.*, et al.* (2013) Epigenetic changes in patients with multiple sclerosis. *Nat Rev Neurol* 9, 35-43

29 Lewén, A.*, et al.* (2000) Free radical pathways in CNS injury. *J Neurotrauma* 17, 871-890

30 Mill, J. and Petronis, A. (2007) Molecular studies of major depressive disorder: The epigenetic perspective. *Mol Psych* 12, 799-814

31 Zigman, W.B. (2013) Atypical aging in down syndrome. *Developmental Disabilities Research Reviews* 18, 51-67

32 Moretti, L.*, et al.* (2012) Cognitive decline in older adults with a history of traumatic brain injury. *Lancet Neurol* 11, 1103-1112

33 Smith, R.L.*, et al.* (2013) Premature and accelerated aging: HIV or HAART? *Frontiers in Genetics* 3, 328

34 Cole, J.H.*, et al.* (2015) Prediction of brain age suggests accelerated atrophy after traumatic brain injury. *Ann Neurol* 77, 571-581

35 Koutsouleris, N.*, et al.* (2013) Accelerated Brain Aging in Schizophrenia and Beyond: A Neuroanatomical Marker of Psychiatric Disorders. *Schizophr Bull* 40, 1140-1153

36 Nenadic, I.*, et al.* (2017) BrainAGE score indicates accelerated brain aging in schizophrenia, but not bipolar disorder. *Psychiatry Research: Neuroimaging* 266, 86-89

37 Pardoe, H.R.*, et al.* (2017) Structural brain changes in medically refractory focal epilepsy resemble premature brain aging. *Epilepsy Research* 133, 28-32

38 Cole, J.H.*, et al.* (2017) Brain-predicted age in Down Syndrome is associated with β-amyloid deposition and cognitive decline. *Neurobiol Aging* 56, 41-49

39 Cole, J.H.*, et al.* (2017) Increased brain-predicted aging in treated HIV disease. *Neurology* 88, 1349-1357

40 Gaser, C.*, et al.* (2013) BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease. *PLoS One* 8

41 Habes, M.*, et al.* (2016) Advanced brain aging: relationship with epidemiologic and genetic risk factors, and overlap with Alzheimer disease atrophy patterns. *Transl Psychiatry* 6, e775

42 Löwe, L.C.*, et al.* (2016) The effect of the APOE genotype on individual BrainAGE in normal aging, Mild cognitive impairment, and Alzheimer's Disease. *PLoS One* 11

43 Ronan, L.*, et al.* (2016) Obesity associated with increased brain age from midlife. *Neurobiol Aging* 47, 63-70

44 Franke, K.*, et al.* (2013) Advanced BrainAGE in older adults with type 2 diabetes mellitus. *Frontiers in Aging Neuroscience* 5, 90

45 Franke, K.*, et al.* (2014) Gender-specific impact of personal health parameters on individual brain aging in cognitively unimpaired elderly subjects. *Frontiers in Aging Neuroscience* 6, 94

46 Habes, M.*, et al.* (2016) White matter hyperintensities and imaging patterns of brain ageing in the general population. *Brain* 139, 1164-1179

47 Steffener, J.*, et al.* (2016) Differences between chronological and brain age are related to education and self-reported physical activity. *Neurobiol Aging* 40, 138-144

48 Luders, E.*, et al.* (2016) Estimating brain age using high-resolution pattern recognition: Younger brains in long-term meditation practitioners. *NeuroImage* 134, 508-513

49 Rogenmoser, L.*, et al.* (2017) Keeping brains young with making music. *Brain Structure & Function*

50 Hayflick, L. (2007) Biological aging is no longer an unsolved problem. *Ann N Y Acad Sci* 1100, 1-13

51 Haufe, S.*, et al.* (2014) On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage* 87, 96-110

52 Fjell, A.M.*, et al.* (2013) Critical ages in the life course of the adult brain: Nonlinear subcortical aging. *Neurobiol Aging* 34, 2239-2247

53 Cherubini, A.*, et al.* (2016) Importance of Multimodal MRI in Characterizing Brain Tissue and Its Potential Application for Individual Age Prediction. *IEEE J Biomed Health Inform* 20, 1232-1239

54 Valizadeh, S.A., *et al.* (2017) Age prediction on the basis of brain anatomical measures. *Hum Brain Mapp* 38, 997-1008

55 Silk, T.J. and Wood, A.G. (2011) Lessons about neurodevelopment from anatomical magnetic resonance imaging. *J Dev Behav Pediatr* 32, 158-168

56 Resnick, S.M., *et al.* (2003) Longitudinal magnetic resonance imaging studies of older adults: A shrinking brain. *J Neurosci* 23, 3295-3301

57 Good, C.D., *et al.* (2001) A Voxel-Based Morphometric Study of Ageing in 465 Normal Adult Human Brains. *NeuroImage* 14, 21-36

58 Blalock, E.M., *et al.* (2003) Gene Microarrays in Hippocampal Aging: Statistical Profiling Identifies Novel Processes Correlated with Cognitive Impairment. *The Journal of Neuroscience* 23, 3807-3819

59 Horn, J.L. and Cattell, R.B. (1967) Age differences in fluid and crystallized intelligence. *Acta Psychologica* 26, 107-129

60 Hedden, T. and Gabrieli, J.D.E. (2004) Insights into the ageing mind: A view from cognitive neuroscience. *Nat Rev Neurosci* 5, 87-96

61 Li, Y., *et al.* (2017) Dependency criterion based brain pathological age estimation of Alzheimer's disease patients with MR scans. *BioMedical Engineering OnLine* 16, 50

Glossary

(411/450 words)

Ageing biomarker: A biological measurement that gives an estimate of an organism's biological age, based on the biological age of an organ, tissue or cell.

Biological age: The hypothetical underlying age of an organism, defined by measuring some aspect of the organism's biology. Biological age may differ from the organism's chronological age and be a better indicator of residual lifespan, functional capacity and risk of age-associated changes.

***Brain age (or brain-predicted age)*:** The predicted age of an individual, derived using high-dimensional neuroimaging data in a machine-learning framework. *Brain age* potentially represents a biomarker of the underlying 'age' of the brain, whereby an 'older' brain in adults indicates increased risks for neurodegenerative diseases and mortality.

Brain ageing: Changes to the human brain that generally accompany ageing. These changes occur at molecular, cellular and tissue levels, and have characteristic functional and behavioural consequences (Box 1).

Deep learning: An extension of machine learning, based on artificial neural networks. 'Deep' refers to the multiple layers of neural networks used, including one or more 'hidden' layers. Each layer is used to transform input data into a different format that encodes something salient about the features contained in the data.

Feature: A variable used in a machine learning algorithm, or an aspect of a dataset that is of some relevance. In the context of *brain age*, features are local measures of brain structure or function (e.g., grey matter volume).

Gerontology: The scientific study of the old and the ageing process.

Machine learning: A statistical approach derived from the study of artificial intelligence, based on the concept that statistical models should be able to make accurate predictions from new 'unseen' data (either categorical, e.g., group membership, or continuous, e.g., age, IQ).

Magnetic Resonance Imaging (MRI): A medical imaging technique that capitalises on the inherent physical properties of biological tissues when inside powerful magnetic fields.

Particularly, hydrogen atoms contained in water within biological tissue behave in characteristic ways when the magnetic fields are manipulated and release energy in the form of radiofrequency (RF) pulses that can be recorded. These RF pulses can be transformed into three-dimensional images that give information on brain volume, blood flow, brain function, white matter microstructure to name but a few biological characteristics.

Voxel: A volume-element, the three-dimensional equivalent of a pixel. Voxels are the unit of resolution for MRI scans of the brain.

Weight maps: Voxel-wise maps of the brain, where each voxel contains a numeric representation of the statistical model learned by a machine learning algorithm.

Text Boxes

Box 1. Brain ageing and its consequences

(292/400 words)

Age-related changes in the human brain are characterised by region-specific and non-linear patterns of highly coordinated and sequenced events of progressive (e.g., cell growth and myelination) and regressive (e.g., synaptic pruning) processes during development [55] and wide-spread atrophy during ageing [56]. In fact, grey matter volume decreases steadily throughout adulthood, while white matter volume follows an inverted 'U-shape' curve, peaking in midlife [56, 57]. Underlying these macroscopic atrophic changes are a whole host of molecular and cellular events. These include altered calcium signalling, genomic alterations, reductions to synaptogenesis and neurite outgrowth, demyelination, microglial activation and subsequent inflammatory responses, changes to cellular metabolism and mitochondrial dysfunction and eventual astrocytic hypertrophy and reduced neuronal activity [58]. These biological changes have behavioural consequences. Most characteristic is the decline in cognitive function commonly observed across adulthood (i.e., cognitive ageing). While memory impairments are most recognised, performance decrements are seen in the majority of cognitive domains, with only crystallised intelligence spared [59]. While the precise relationship between cognitive ageing and the neurophysiological changes in the brain is still unclear, the presence of some link between the two is intuitive [60]. Beyond cognitive ageing, advanced brain ageing is also associated with an increased prevalence of brain diseases, particularly neurodegenerative diseases, including Alzheimer's, Parkinson's and Amyotrophic Lateral Sclerosis, to name a few. In fact, age is the biggest risk factor for many of these diseases, and the progressive nature of these conditions means that severity worsens with age. The dementia that results from many of these diseases causes a high burden on society and on individuals, both financially and socially. Currently, there are limited options for modifying or treating these diseases, and even evaluating potential therapeutics is difficult as the relatively slow rates of disease progression make long-term interventional studies challenging.

Box 2. How brain age prediction works

(396/400 words)

The general analytic 'pipeline' for predicting the biological age of individual brains uses structural

neuroimaging from a large sample of healthy people, screened to exclude those with neuropsychiatric or physical health conditions. The chronological age of these individuals is known and they should represent the adult lifespan. These individuals comprise the so-called 'training set'. The following stages are carried out: 1) Neuroimaging data from the training set then usually undergoes image pre-processing to derive meaningful features that relate to ageing, for example spatial registration to a template, to quantify brain volume at each voxel. 2) These features are then used as predictors or independent variables in a regression model, with chronological age as the outcome or dependent variable. This is the inverse to conventional statistical approaches that aim to understand which brain regions may have a linear relationship with age, as in voxel-based morphometry. Ordinary least-squares (OLS) regression models are inappropriate for such high-dimensional neuroimaging datasets, where each individual is characterized by several hundred or thousand datapoints. Hence, multivariate machine-learning methods (e.g., support vector, relevance vector or Gaussian Processes regression) are used, as they were designed to cope with high-dimensional types of data. 3) The accuracy of the machine learning regression model is assessed using a cross-validation procedure. Popular variations of this include k-fold and split-half cross-validation. The idea behind cross-validation is that some proportion of the individuals in the training set is left out of the initial 'learning' stage. The parameters of the learned model (analogous to OLS beta estimates) are then applied to the pre-processed data of the left-out individuals, resulting in brain-derived predictions of age. This age prediction is then compared to the known chronological age of each left-out individual. Accuracy metrics, including the Pearson correlation between the predicted and chronological ages, the R^2 (i.e., variance explained) of the prediction model and the mean absolute error (MAE), are then generated to evaluate the specific age prediction model. 4) Assuming the *brain age* prediction model reaches a desired level of accuracy, entirely new individuals ('test set') can now be run through the model, generating individual predictions of *brain age*. The difference between predicted and chronological age quantifies the acceleration or deceleration in individual brain ageing. For example, if the *brain age* of a 70 year-old results in a difference of +5 years, this individual shows the typical atrophy pattern of a 75 years old.

Figure Legends

Figure 1. Trajectories of biological ageing

As chronological age increases, there is a trend towards a higher risk of diseases and the onset of cognitive decline. This trend is thought to have a biological basis, relating to the cumulative damage to cells and tissues acquired over time. While people who are generally healthy (grey arrow) reach the threshold for symptoms to appear at approximately a similar age, other people may follow different trajectories of biological ageing. This could be due to genetic differences, or exposure to environmental effects that subtly increase the rate at which age-associated damage accumulates (blue arrow). Potentially, people may experience pronounced environmental influences, such as a brain injury or cerebral infection, leading to a marked acceleration of the rate of biological ageing (red line). In the current context, *brain age* may represent a measure of the underlying biological age of the brain. By measuring how far an individual is from the healthy brain ageing trajectory, researchers hope to be able to quantify advanced and decelerated brain ageing and use this to predict individual's future trajectories and subsequent risk of age-associated health deterioration.

Figure 2. How brain age prediction works

Overview of the *brain age* prediction process, using 'supervised' machine learning. A) Neuroimaging data, usually T1-weighted structural MRI scans, from healthy individuals (training set) are labelled with the participants' chronological age and put into a machine learning regression model. B) To validate the accuracy of the model, a proportion of participants' images are left out of the model. For example, ten-fold cross-validation involves training the model on 90% of participants and predicting age values on the left-out 10%. This is then iterated through all participants, and predicted values are compared with real values (i.e., chronological age) to assess accuracy. C) Assuming the model is sufficiently accurate, the model is trained using the entire training set and the resulting model coefficients applied to new participants' brain scans (test set) to generate unbiased individual *brain age* predictions, in this example 61.7 years. D) The predicted *brain age* can then be compared with the chronological age of test set participants, with 'older' appearing brains assumed to reflect advanced brain ageing and 'younger' appearing brains to reflect decelerated or healthy brain ageing. The discrepancy between *brain age* and chronological age (brain-predicted age difference) can then be used as a metric to statistically

relate to other measured characteristics of the participants.

Acknowledgements

The authors would like to thanks Drs Rob Leech, Romy Lorenz and Ines Violante for their insightful and instructive comments on drafts of this article. This work was supported by the German Research Foundation [DFG; Project FR 3709/1-1 to KF]. The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

501 # Tables

502 ## Table 1. Studies assessing *brain age* in neurological and psychiatric diseases

Clinical group	N	Age mean (SD)	Features for *brain age*	Mean *brain age* difference (years)	References
Alzheimer's disease	102	76 (8)	GM	10.0	Franke, 2010 [8]
Alzheimer's disease	150	75 (8)	GM	baseline: 6.7 follow-up (2y): 9.0	Franke, 2012 [13]
Alzheimer's disease (APOE ε4 carriers/non-carriers)	101 / 49	74 (7) / 76 (9)	GM	baseline: 5.8 / 6.2 follow-up (2y): 8.3 / 7.7	Löwe, 2016 [42]
Alzheimer's disease	411	75 (7)	Hippocampus	7	Li, 2017 [61]
At-risk mental states for psychosis	89	25 (6)	GM	1.7	Koutsouleris, 2013 [35]
Bipolar disorder	22	38 (11)	GM	-1.3 (males: -1.9 / females: -0.8)	Nenadic, 2017 [36]
Borderline personality disorder	57	26 (7)	GM	3.1	Koutsouleris, 2013 [35]
Diabetes mellitus type 2	98	65 (8)	GM	4.6	Franke, 2013 [44]
Diabetes mellitus type 2	12	63 (7)	GM	baseline: 5.1 follow-up (4y): 5.9	Franke, 2013 [44]
Down's syndrome	46	42 (9)	Whole brain	2.5	Cole, 2017 [38]
Epilepsy (medically-refractory/newly-diagnosed)	94 / 42	32 (14) / 31 (11)	Whole brain	4.5 / 0.9	Pardoe, 2017 [37]
HIV	162	56	Whole brain	2.2	Cole, 2017 [39]
Major depression	104	42 (8)	GM	4.0	Koutsouleris, 2013 [35]
Mild cognitive impairment, progressive	112	74 (7)	GM	baseline: 6.2 follow-up (3y): 9.0	Franke, 2012 [13]

Mild cognitive impairment, progressive (early/late)	58 / 75	74 (7) / 75 (7)	GM	8.7 / 5.6	Gaser, 2013 [40]
Mild cognitive impairment, progressive (APOE ε4 carriers / non-carriers)	78 / 34	74 (6) / 75 (9)	GM	baseline: 5.8 / 5.5 follow-up (3y): 8.7 / 7.3	Löwe, 2016 [42]
Mild cognitive impairment, stable	36	77 (6)	GM	baseline: -0.5 follow-up (3y): -0.4	Franke, 2012 [13]
Mild cognitive impairment, stable (APOE ε4 carriers/non-carriers)	14 / 22	77 (6) / 77 (6)	GM	baseline: -0.9 / -0.9 follow-up (3y): 0.0 / -0.6	Löwe, 2016 [42]
Obesity	227	58 (17)	WM	10	Ronan, 2016 [43]
Objective cognitive impairment (mild/major)	632 / 251	58 (15) / 58 (16)	Whole brain (multimodal)	0.7 / 1.7	Liem, 2017 [10]
Schizophrenia	141	28 (12)	GM	5.5	Koutsouleris, 2013 [35]
Schizophrenia	341	34 (12)	GM	baseline: 3.4 follow-up (4y): 4.3	Schnack, 2016 [12]
Schizophrenia	45	34 (10)	GM	2.6 (males: 3.4 / females: 1.1)	Nenadic, 2017 [36]
Traumatic brain injury	99	38 (12)	GM / WM	4.7 / 6.0	Cole, 2015 [34]

Features for *brain age* are reported as the aspect of brain structure used predictors in the *brain-age* model. GM = grey matter, WM = white matter. Data formats include voxelwise 3D images or summary measures of cortical thickness and subcortical volumes. Multi-modal refers to a combination of structural and function MRI.